Least Squares Estimation Methods

Ordinary Least Squares (OLS)

Generalized Least Squares (GLS)

Feasible Generalized Least Squares (FGLS)

- Weighted Least Squares (WLS)

  Heteroscedasticity

- Cochrane-Orcutt Technique

  Autocorrelation

OLS Matrix Approach

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i \quad \text{where } i = 1, 2, 3, \ldots, n$$

$$Y_1 = \beta_1 + \beta_2 X_{21} + \beta_3 X_{31} + \cdots + \beta_k X_{k1} + u_1$$

$$Y_2 = \beta_1 + \beta_2 X_{22} + \beta_3 X_{32} + \cdots + \beta_k X_{k2} + u_2$$

$$\vdots$$

$$Y_n = \beta_1 + \beta_2 X_{2n} + \beta_3 X_{3n} + \cdots + \beta_k X_{kn} + u_n$$

$$\underset{n\times 1}{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}} = \underset{n\times k}{\begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix}} \cdot \underset{k\times 1}{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}} + \underset{n\times 1}{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}}$$

Matrix Notation $\underset{n\times 1}{y} = \underset{n\times k}{X}\ \underset{k\times 1}{\beta} + \underset{n\times 1}{u}$

OLS Matrix Approach

Matrix Notation $\underset{n\times 1}{y} = \underset{n\times k}{X}\,\underset{k\times 1}{\beta} + \underset{n\times 1}{u}$

$$u = \underset{n\times 1}{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}} \;\&\; u' = \underset{1\times n}{\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}},\; \hat{u}'\hat{u} = \sum_{i=1}^{n} \hat{u}_i^2$$

$$\hat{u}'\hat{u} = \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right) = y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}$$

$$\frac{\partial \hat{u}'\hat{u}}{\partial \hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

$$2X'X\hat{\beta} = 2X'y$$

$$\hat{\beta} = \left(X'X\right)^{-1}X'y$$

Variance-Covariance Matrix

Assume normal distribution: $u \sim N(0, \sigma^2 I)$

Under OLS Assumptions

$$\sum_{n\times n} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_{\mathrm{n}}^2 \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 I$$

Under Heteroscedasticity Problem

$$\sum_{n\times n} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

Weighted Least Squares (WLS)

Scalar

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$\frac{Y_i}{\sigma_i} = \beta_1\left(\frac{1}{\sigma_i}\right) + \beta_2\left(\frac{X_i}{\sigma_i}\right) + \left(\frac{u_i}{\sigma_i}\right)$$

$$Y_i^* = \beta_1^* X_{0i}^* + \beta_2^* X_i^* + u_i^*$$

Matrix

$$\underset{k\times 1}{\hat{\beta}} = (\underbrace{\underset{k\times n}{X'}\ \underset{n\times n}{\Sigma^{-1}}\ \underset{n\times k}{X}}_{k\times k})^{-1}\ \underbrace{\underset{k\times n}{X'}\ \underset{n\times n}{\Sigma^{-1}}\ \underset{n\times 1}{y}}_{k\times 1}$$

Autocorrelation: Cochrane-Orcutt

Scalar $$Y_t = \beta_1 + \beta_2 X_t + u_t$$

$$u_t = \rho u_{t-1} + \varepsilon_t$$

1. Estimate model using OLS and obtain estimated u_t $\quad \hat{u}_t = \hat{\rho}\hat{u}_{t-1} + \upsilon_t$
2. Compute
3. $\left(Y_t - \hat{\rho} Y_{t-1}\right) = \beta_1 \left(1 - \hat{\rho}\right) + \beta_2 \left(X_t - \hat{\rho} X_{t-1}\right) + \varepsilon_t$

$$Y_t^* = \beta_1^* + \beta_2^* X_t^* + \varepsilon_t^*$$

4. Iterative procedure to estimate ρ

Autocorrelation: Cochrane-Orcutt

Variance-Covariance Matrix $\underset{n\times n}{\Sigma} = \sigma^2 \underset{n\times n}{\Omega}$

where

$$\underset{\mathrm{n\times n}}{\Omega} = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \rho^2 & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \rho & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & \rho & 1 \end{bmatrix}$$

Matrix

$$\underset{k\times 1}{\hat{\beta}} = (\underset{k\times n}{X'} \underset{n\times n}{\hat{\Omega}^{-1}} \underset{n\times k}{X})^{-1}_{k\times k} \underset{k\times n}{X'} \underset{n\times n}{\hat{\Omega}^{-1}} \underset{n\times 1}{y} \quad (k\times 1)$$

OLS vs GLS vs FGLS

OLS Estimation

$$\underset{n\times n}{\Sigma} = \sigma^2 \underset{n\times n}{I}$$

$$\underset{k\times 1}{\hat{\beta}} = (\underset{k\times k}{X'X})^{-1} \underset{k\times n}{X'} \underset{n\times 1}{y}$$

WLS Estimation

$$\underset{n\times n}{\Sigma} = \hat{\sigma}_i^2 \underset{n\times n}{I}$$

$$\underset{k\times 1}{\hat{\beta}} = (\underbrace{\underset{k\times n}{X'} \underset{n\times n}{\hat{\Sigma}^{-1}} \underset{n\times k}{X}}_{k\times k})^{-1} \underbrace{\underset{k\times n}{X'} \underset{n\times n}{\hat{\Sigma}^{-1}} \underset{n\times 1}{y}}_{k\times 1}$$

Cochrane-Orcutt

$$\underset{n\times n}{\Sigma} = \sigma^2 \underset{n\times n}{\hat{\Omega}}$$

$$\underset{k\times 1}{\hat{\beta}} = (\underbrace{\underset{k\times n}{X'} \underset{n\times n}{\hat{\Omega}^{-1}} \underset{n\times k}{X}}_{k\times k})^{-1} \underbrace{\underset{k\times n}{X'} \underset{n\times n}{\hat{\Omega}^{-1}} \underset{n\times 1}{y}}_{k\times 1}$$

FGLS Estimation

$\underset{n\times n}{\Sigma}$ *is not known*

$$\underset{k\times 1}{\hat{\beta}} = (\underbrace{\underset{k\times n}{X'} \underset{n\times n}{\hat{\Sigma}^{-1}} \underset{n\times k}{X}}_{k\times k})^{-1} \underbrace{\underset{k\times n}{X'} \underset{n\times n}{\hat{\Sigma}^{-1}} \underset{n\times 1}{y}}_{k\times 1}$$

GLS Estimation

$\underset{n\times n}{\Sigma}$ *is known*

$$\underset{k\times 1}{\hat{\beta}} = (\underbrace{\underset{k\times n}{X'} \underset{n\times n}{\Sigma^{-1}} \underset{n\times k}{X}}_{k\times k})^{-1} \underbrace{\underset{k\times n}{X'} \underset{n\times n}{\Sigma^{-1}} \underset{n\times 1}{y}}_{k\times 1}$$

Dummy Variable

In the case that we have special event that cannot be quantify (i.e. economics crisis), dummy variable can be used.

Dummy variable is a discrete and binary-choice variable (0 or 1) that used to quantify qualitative independent variable.

Dummy variable can only be 0 or 1. It cannot be 1 or 2.

In case that qualitative variable has more than 2 choices, the model will have more than 1 dummy variable.

(Two choice – 1 dummy), (Three choice – 2 dummy), (Four choice – 3 dummy).

Intercept Dummy Variable

General model $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$

Model with Intercept Dummy Variable

$$Y_t = \beta_0 + \gamma_0 D_t + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$$

where: $D_t = 0$ before crisis

$= 1$ after crisis.

This model can be interpreted as:

Before Crisis: $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$

After Crisis: $Y_t = (\beta_0 + \gamma_0) + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$

Intercept & Slope Dummy Variables

Model with Intercept and Slope Dummy Variable

$$Y_t = \beta_0 + \gamma_0 D_t + \beta_1 X_{1t} + \gamma_1 D_t X_{1t} + \beta_2 X_{2t} + \gamma_2 D_t X_{2t} + u_t$$

where: $D_t = 0$ before crisis
$= 1$ after crisis.

This model can be interpreted as:

Before Crisis: $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$

After Crisis: $Y_t = (\beta_0 + \gamma_0) + (\beta_1 + \gamma_1) X_{1t} + (\beta_2 + \gamma_2) X_{2t} + u_t$

Test Structural Change – Dummy Variable Test vs Chow Test

Dummy Variable Technique

R	All	$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$
UR	Before	$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$
	After	$Y_t = (\beta_0 + \gamma_0) + (\beta_1 + \gamma_1) X_{1t} + (\beta_2 + \gamma_2) X_{2t} + u_t$

Chow Test

1	All	$Y_t = \lambda_0 + \lambda_1 X_{1t} + \lambda_2 X_{2t} + u_t$	*for t = 1, 2, …,* $n_1 + n_2$
2	Before	$Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + u_{1t}$	*for t = 1, 2, …,* n_1
3	After	$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_{2t}$	*for* $t = n_1+1, n_1+2, \ldots, n_1+n_2$

Dummy Variable vs Chow Test

Chow Test

$$H_0: \quad \alpha_0 = \beta_0 = \lambda_0$$
$$\text{and} \quad \alpha_1 = \beta_1 = \lambda_1$$
$$\text{and} \quad \alpha_2 = \beta_2 = \lambda_2$$
$$H_a: \quad \text{Otherwise}$$

$$F = \frac{(RSS_1 - RSS_2 - RSS_3)/k}{(RSS_2 + RSS_3)/(n_1 + n_2 - 2k)}$$

Dummy Variable

$$H_0: \quad \gamma_0 = \gamma_1 = \gamma_2 = 0$$
$$H_a: \quad \text{Otherwise}$$

$$F = \frac{(RSS_R - RSS_{UR})/k}{RSS_{UR}/(n-k)}$$

Comparing the Two Methods

Chow test is to test whether all estimators are the same or not.

Dummy variables technique test whether intercept dummy coefficient and all slope dummy coefficients are all equal to zero or not.

Interaction Effects Using Dummy Variables

Model with Intercept and Interaction Dummy Variables

$$Y_t = \alpha_0 + \alpha_1 D_{Mt} + \alpha_2 D_{Jt} + \alpha_3 D_{Mt} D_{Jt} + \beta X_t + u_t$$

where: $D_{Mt} = 0$ other days or $= 1$ Monday.

$D_{Jt} = 0$ other months or $= 1$ January.

The coefficients of the dummy variables can be interpreted as:

α_1 = Monday effect.

α_2 = January effect.

α_3 = Monday in January effect.